

5.2
409/2:12
C.2

NORTH CAROLINA STATE LIBRARY
RALEIGH

N. C.
Doc.

PHSB STUDIES

OCT 23 1978

A Special Report Series by the N.C. Department of Human Resources, Division of Health Services, Public Health Statistics Branch, P.O. Box 2091, Raleigh, N.C.

No. 12

October 1978

SMALL AREA STATISTICS: PROBLEMS WITH VITAL STATISTICS RATES

For many years—since 1932—the Division of Health Services has published resident birth and death rates for North Carolina counties and major cities. While aware of certain inconsistencies between recording of residence on birth and death certificates and Census enumerations of resident populations, we regarded the magnitude of these inconsistencies as minimal and have proceeded to do the job we felt charged to do, that is, to generate and disseminate data presumably helpful to health officials and researchers in identifying the particular health problems of cities, counties and various regions of the state.

Beginning with 1975 vital statistics reports, however, we had come to recognize that annexations generally preclude reliable intercensal city population estimates in the required detail, and we discontinued the computation of birth and death rates. At the same time, we felt reasonably secure in the use of census data and corresponding vital event data and early this year computed 1968-72 age-race-sex-adjusted death rates (cause-specific) for 38 cities, those that were incorporated and exceeded 10,000 population in 1970. Findings were reported in a PHSB study entitled "Mortality in North Carolina Cities" (1) after satisfying ourselves that data counts were correct (insofar as original sources were correct) and that adjustment procedures were accurately programmed. Also at that time, we investigated whether Morganton's low death rate (lowest of the 38 cities) might be due to the inclusion of Broughton Hospital residents in rate denominators. According to Rand McNally (2) and later, the Bureau of the Census (personal communication), Morganton's 1970 population of 13,625 excluded Broughton Hospital which was reported "outside corporate limits." Thus, we published these data, providing—for the first time ever—comparisons of mortality levels among cities on a cause-specific basis. These data, which showed wide variation in age-race-sex-adjusted city mortality levels, were widely publicized by the news media and generated considerable interest and concern among health officials and researchers.

Detection of Problems: The Morganton Case

Contrary to earlier information, we are now informed by the Bureau of the Census that Morganton's 1970 population count did include Broughton Hospital residents in accordance with the city map for 1970. This map did not identify Broughton Hospital as a "political island" as was true of the map used in the 1960 Census. Hence, since deaths to residents of long-term health and penal institutions are by state regulation allocated to the decedent's place of residence prior to admission (3), Morganton's "population at risk" (denominator used in rate computation) was substantially inflated, Broughton Hospital accounting for about 2,100 of Morganton's 13,625 residents in 1970. This contributed greatly to Morganton's having the lowest of the 38 city death rates.

All other things being equal, other cities with sizable inmate populations within their corporate boundaries similarly would have artificially low death rates. Among the 38 cities, Salisbury and Raleigh appear particularly at risk, based on the

institution populations enumerated in 1970 (see table on page 3). And in fact, it is noteworthy that in the referenced study (1), Raleigh had the third and Salisbury the fourth lowest of the 38 city rates.

Detection of Problems: The Hickory Case

At the opposite end of the death rate spectrum, Hickory had the highest of the 38 city rates (age-race-sex-adjusted). Given the Morganton situation, we proceeded to investigate the numerators and denominators used to compute Hickory's rates and, unfortunately, errors were found.

Here, errors involved inflated death counts resulting from the inclusion of decedents who had lived outside Hickory corporate limits, for example, in Longview, Brookford, and other outlying areas having Hickory mailing addresses. The PHSB supplied the Hickory Engineering Department a sample of 100 Hickory addresses that were "inside city limits" according to death certificates. An audit by the Engineering Department revealed that 17 of these addresses were definitely outside Hickory corporate limits. Eight of those 17 decedents were from Longview which is an incorporated place, so the informant was correct in saying the decedent's residence was "inside city limits." However, the Hickory address was misleading; coders of death certificates were lead to believe these decedents lived inside Hickory city limits.

Other cities with well-populated areas outside corporate limits may have similar problems.

Misleading Death Rates

Problems with Morganton denominators and Hickory numerators probably involve all cities to some degree. Certainly, all researchers should be aware of the following practices with respect to residence allocation:

- (i) Presently, both the Census and vital event resident rules consider military personnel and college students as residents of the military or college area (3,4). In practice, however, a permanent residence (e.g., the parental address) is sometimes given on student death certificates and often given on military death certificates as revealed by an audit of deaths occurring in Orange and Cumberland counties.
- (ii) The Census enumerates inmates of long-term health and penal institutions as residents of the area where the institution is located (4). However, North Carolina rules on death certification (3) specify that usual residence prior to admission be given for deaths in hospitals, sanatoria and all other institutions (mental, penal, old age, orphan, nursing home, rest home, etc.).

Regarding (i) above, since relatively few college students die, that situation usually would not greatly affect death rates that are adjusted for age. However, young military personnel are at a greater risk of death, and that situation may appreciably lower death rates for military bases and corresponding counties. Also, for studies of relatively infrequent events, for example, a study of suicide among young people, both situations represent a potential bias.

Population in Group Quarters
38 North Carolina Cities, 1970

<u>City</u>	<u>Total Population</u> ¹	<u>Percent in Institutions</u> ²	<u>Percent in Other Group Quarters</u> ³
Albemarle	11,126	0.4	0.8
Asheboro	10,797	0.6	0.2
Asheville	57,681	1.3	1.2
Burlington	35,930	0.0	0.3
Chapel Hill	25,537	0.1	29.6
Charlotte	241,178	0.5	1.4
Concord	18,464	0.4	3.3
Durham	95,438	0.6	7.5
Eden	15,871	0.2	0.4
Elizabeth City	14,069	0.6	0.9
Fayetteville	53,510	1.7	1.7
Gastonia	47,142	0.6	0.2
Goldsboro	26,810	0.3	0.6
Greensboro	144,076	0.7	5.5
Greenville	29,036	0.5	14.4
Henderson	13,896	1.0	0.5
Hickory	20,569	0.0	4.0
High Point	63,204	0.7	1.7
Jacksonville	16,021	0.1	0.1
Kinston	22,309	0.3	0.5
Lenoir	14,705	0.5	0.2
Lexington	17,205	0.8	0.1
Lumberton	16,961	0.4	0.2
Monroe	11,282	1.0	0.4
Morganton	13,625	16.4	5.2
New Bern	14,660	0.4	0.3
Raleigh	121,577	4.1	7.3
Reidsville	13,636	0.1	0.4
Roanoke Rapids	13,508	0.4	0.8
Rocky Mount	34,284	0.2	0.3
Salisbury	22,515	4.8	6.1
Sanford	11,716	0.2	0.5
Shelby	16,328	0.1	0.1
Statesville	19,996	0.7	0.3
Thomasville	15,230	2.4	0.4
Wilmington	46,169	0.3	0.3
Wilson	29,347	1.9	2.8
Winston Salem	132,913	0.8	3.3

¹Uncorrected Census count, April 1, 1970 (4).

²Includes homes, schools, hospitals or wards for juveniles and the physically or mentally handicapped; hospitals for mental, tuberculosis or chronic disease patients; homes for unwed mothers; nursing, convalescent and rest homes for the aged and dependent; and correctional institutions (4).

³Includes rooming and boarding houses, communes, farm and nonfarm workers' dormitories, convents, monasteries, military barracks, ships, college dormitories, sorority and fraternity houses, institutional quarters for staff, missions, flophouses, Salvation Army shelters, etc. (4).

Regarding (ii), places with large health care and penal populations will definitely have artificially low death rates, given that deaths are allocated according to current regulations.

The table on page 3 shows each city's 1970 population and the percentages enumerated in institutions and other group quarters.

Misleading Birth and Fertility Rates

Given the preceding conventions for Census enumerations, it is easily seen that places with large institutional populations will have low birth and fertility rates relative to their true "at risk" population. Although counted in the denominators, institutionalized persons and persons in certain other group quarters are at relatively low risk of contributing to the numerator of a birth or fertility rate. On this basis, one may prefer that denominators be restricted to the population in households (which excludes the population in group quarters). However, these population bases are not available intercensally in North Carolina. Also, use of the household population data would mean that the numerator of a rate could include events not represented in the denominator, for example, a decedent in an institution would not be included in any city or county denominator. Additionally, a county or city could no longer compare itself to places outside North Carolina since other places would be computing rates based on a different kind of denominator.

The Need for City Statistics

The PHSB routinely publishes reports based on age-race-sex-adjusted death rates. These are and have been of great use to health planners at the state, regional and county levels. In addition to these reports, the PHSB and other users of vital statistics have used these rates for a variety of research projects, often for ecologic studies that attempt to find associations between disease-specific county rates and certain environmental factors (measures of pesticide use, water quality, rurality, etc.).

Ecologic studies are to be distinguished from the more rigorous, definitive, and expensive epidemiologic studies that are based on individuals as the sampling units with individuals being scrutinized for disease classification, exposure histories and other relevant information. The sampling unit of an ecologic study is a set or a defined classification of people; typically, the units are geographically-based, usually counties. Measures of health status (often mortality rates) are collected for each county. In addition, exposure determinations are made for the counties. Even though there may be an observed relationship between the rates and the exposure classifications over the counties, a lack of specificity of what is operative within counties regarding individuals (i) precludes estimates of individual risks and (ii) leaves us blind to possible spurious results (ecologic fallacies).

Environmental factors may vary widely within a county; these factors relate to the number of towns, sizes of towns, industrial and agricultural composition to name a few. Water quality is one important factor and will be used to illustrate some potential problems in the use of small area statistics.

It is almost axiomatic that water sources and water constituent values often fluctuate radically over the geography of a county. This means that to classify counties by several water variables is either a formidable task relying on many sources of data or it leads to unreliable results.

There are quite a few studies of relationships between health and water where the unit of study is a county. Some of these studies use elaborate procedures to estimate water quality on the county level. Nevertheless, there remains the disconcerting fact that we are left wide open for a classical ecological fallacy, a gap that is to some degree narrowed as we sample smaller subareas. A city's water characterizations may be comprised of less contrasting parts than would be true at the county level; so a natural and perhaps necessary step is to consider small and, in terms of exposure (water constituents), more homogeneous areas for the sampling units. This is an important consideration, especially when one perceives that the water characteristics of a county—which contains several towns, rivers, industries, etc.—may be so mixed and complex that any water classification would be at best a symbolic descriptor composed of several variables and would vaguely mimic what we are really after.

But to go with smaller areas such as cities, which is the only other possibility using existing data, is not without its problems, as has been depicted in the earlier sections of this report. New types of classification problems now emerge and the question is: are these any simpler to solve? They may be. But they must be recognized first. The types of problems encountered in using cities rather than counties pertain to the placing of individuals who are ultimately counted in the numerator and/or denominator of a rate. That is, rates for cities are to a greater extent perturbed by errors and disparities in vital records and census reports concerning place of residence. Such errors are not only more likely to occur in collecting city data than in collecting county data (e.g., errors often involve city boundaries) but when they do occur they are known to take a greater toll on city rates than on county rates. To summarize, as the geographic area becomes smaller the relative magnitude of the data quality problem tends to shift from concerns involving the ascertainment of exposure (independent variable) to those of determining the rates (dependent variable).

So far we have mentioned a few considerations that should enter into a decision of whether to work with county rates or city rates. These are contrasted as follows:

- The exposure factors that correspond to city rates are easier to define and measure than those of county rates.
- The boundary lines for cities come closer to distinguishing real environmental differences than do county lines.
- The populations on which the two kinds of rates are based are different. Cities are subsets of counties and, except under certain conditions, city rates do not extend to the state population as a whole but rather to the state's urban population. A comparison of city and county rates must take this into account.
- City rates are most accurate on census years. Demographic projections to intercensal years are probably more accurate for counties than for cities.

- The numerators and denominators are much greater for county rates. This is particularly important for rare diseases. City rates are more sensitive to counting errors due to mistaking residences on death certificates and to differences between the numerators and denominators regarding residence status of persons in certain large institutions.

There are other contrasts but these seem to be some essential ones. The first two favor the use of city rates and the remaining ones tend to favor county rates. While investigators recognize most of these, it appears that the first item mentioned in the list is often unrecognized or else its importance is considered secondary to some of the others.

A statistical comparative study is currently underway to determine the extent of the differences between city and county adjusted mortality rates and to inquire as to what factors determine such differences.

Where from Here?

According to the Division of Vital Statistics, National Center for Health Statistics, that agency has for several years recommended that states collect residence information that is completely consistent with Census enumeration practices. To date, however, this has not been done in North Carolina or any of several other states investigated; and to do so will amount to a tremendous effort in terms of re-educating registrars, funeral directors, physicians, hospitals and others involved in the registration of vital events. Also, this will not solve problems inherent in traditional birth and fertility rates or death rates unadjusted for age; denominators will still overrepresent the "true" population at risk.

Thus, there presently seems no immediate solution to the problems of small area statistics, and we can only suggest the following:

- Persons using population-based city and county rates of any kind should be acutely aware of possible inconsistencies between the Census and their particular data. It is recommended that city rates not be used to study rare events at this time.
- Health Departments (registrars) serving cities with well-populated areas outside corporate limits should set in place some mechanism for assuring that city vital event counts exclude residents of those outlying areas; else, counts of births, deaths, etc. will be inflated.
- The Vital Records Branch (VRB) of the Division of Health Services and local registrars should work to insure that funeral directors and hospitals understand and abide by state regulations with respect to residence allocation.
- The VRB should receive notification from the Office of State Budget and Management and inform local registrars and funeral directors of city annexations in order that city counts of vital events be more accurate.

- Due to annexation and late corrections in Census counts, a city's intercensal population estimates may be unreliable in terms of age, race and sex distributions; thus, intercensal vital event rates should not be computed for cities, especially those that have annexed postcensally.
- The PHSB and the Office of State Budget and Management should prepare and regularly update a comprehensive inventory of institutional beds in the various cities and counties in order to keep abreast of corresponding impacts on population characteristics and birth and death rates.
- The PHSB and the VRB should investigate the feasibility of collecting residence information that is completely consistent with Census enumeration practices. To do so will make for more accurate death rates in cases of large institutional populations but will make North Carolina unlike many other states.

To end on a more optimistic note, two future developments could enhance the accuracy of small area statistics. One, the Postal Service has recently announced a new 9-digit zip code system that will allow mail to be sorted down to individual street blocks (5). If birth and death certificates required these place-of-residence zip codes, we could identify non-city residents as well as residents of particular city segments. The problem here is whether the 9-digit zip system will be accepted to the extent that informants can supply the required number.

The second and probably more certain development is implementation of a quinquennial census after 1980. Then, given reliable census counts, projections for intercensal years should be more accurate than those presently based on a decennial census.

Together, the above two developments offer new potential for small area statistics. Although, for reasons of confidentiality, we would not produce vital statistics for city blocks, we could produce—every five years at least—rates for city segments such as census tracts or zip-code areas detailed to 7 or 8 digits. Thus might we finally acquire those tools envisioned by our forebears and described in the following passage from the North Carolina Board of Health's Biennial Report of 1909-1910:

To the health officer, vital statistics are what the chart and compass are to the mariner. Without them he knows not whence he came nor whither he goes.

REFERENCES

- (1) North Carolina Department of Human Resources, Division of Health Services, Public Health Statistics Branch. "Mortality in North Carolina Cities," PHSB Studies. No. 9. Raleigh, February 1978.
- (2) Rand McNally Standard Reference Map and Guide of North Carolina. Rand McNally and Company, Chicago, 1972.
- (3) North Carolina Department of Human Resources, Division of Health Services. "Funeral Directors' Handbook on Death and Fetal Death Registration." Raleigh, December 1975.
- (4) U.S. Bureau of the Census. Census of Population: 1970 General Population Characteristics Final Report PC(1)-B35 North Carolina. U.S. Government Printing Office, Washington, D.C., 1971.
- (5) "Zip: Longer to Write, Faster to Mail," Raleigh (N.C.) News and Observer, September 14, 1978, p. 1.



3 3091 00747 2467

Public Health Statistics Branch
Division of Health Services
Department of Human Resources
P.O. Box 2091
Raleigh, North Carolina 27602